#### SEQUENCE LISTING

- (1) GENERAL INFORMATION:
  - (i) APPLICANTS: Seed, Brian et al.
  - (ii) TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected Cells by Chimeric CD4 Receptor-Bearing Cells
  - (iii) NUMBER OF SEQUENCES: 27
    - (iv) CORRESPONDENCE ADDRESS:
      - (A) ADDRESSEE: Fish & Richardson
      - (B) STREET: 225 Franklin Street
      - (C) CITY: Boston
      - (D) STATE: MA
      - (E) COUNTRY: USA
      - (F) ZIP: 02110-2804
      - (v) COMPUTER READABLE FORM:
        - (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
        - (B) COMPUTER: IBM PS/2 Model 50Z or 55SX
        - (C) OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
        - (D) SOFTWARE: Wordperfect (Version 5.0)
    - (vi) CURRENT APPLICATION DATA:
      - (A) APPLICATION NUMBER:
      - (B) FILING DATE:
      - (C) CLASSIFICATION:
    - (vii) PRIOR APPLICATION DATA:
      - (A) APPLICATION NUMBER: 07/847,566
      - (B) FILING DATE: March 6, 1992
      - (C) CLASSIFICATION:
    - (vii) PRIOR APPLICATION DATA:
      - (A) APPLICATION NUMBER: 07/665,961
      - (B) FILING DATE: March 7, 1991
      - (C) CLASSIFICATION:
  - (viii) ATTORNEY/AGENT INFORMATION:
    - (A) NAME: Clark, Paul T.
    - (B) REGISTRATION NUMBER: 30,162
    - (C) REFERENCE/DOCKET NUMBER: 00786/212001
    - (ix) TELECOMMUNICATION INFORMATION:
      - (A) TELEPHONE: (617) 542-5070
      - (B) TELEFAX: (617) 542-8906
        - (C) TELEX: 200154
  - (2) INFORMATION FOR SEQ ID NO:1:
    - (i) SEQUENCE CHARACTERISTICS:
      - (A) LENGTH: 1728 base pairs

region in the sign of the property of the second of the se

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

# (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC	50
GCTCCTCCCA GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG	100
GGGATACAGT GGAACTGACC TGTACAGCTT CCCAGAAGAA GAGCATACAA	150
TTCCACTGGA AAAACTCCAA CCAGATAAAG ATTCTGGGAA ATCAGGGCTC	200
CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT GACTCAAGAA	250
GAAGCCTTTG GGACCAAGGA AACTTCCCCC TGATCATCAA GAATCTTAAG	300
ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA	350
GGTGCAATTG CTAGTGTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC	400
TTCAGGGGCA GAGCCTGACC CTGACCTTGG AGAGCCCCCC TGGTAGTAGC	450
CCCTCAGTGC AATGTAGGAG TCCAAGGGGT AAAAACATAC AGGGGGGGGAA	500
GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC ACCTGGACAT	550
GCACTGTCTT GCAGAACCAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG	600
GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGA	650
ACAGGTGGAG TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG	700
GCAGTGGCGA GCTGTGGTGG CAGGCGGAGA GGGCTTCCTC CTCCAAGTCT	750
TGGATCACCT TTGACCTGAA GAACAAGGAA GTGTCTGTAA AACGGGTTAC	800
CCAGGACCCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC CACCTCACCC	850
TGCCCCAGGC CTTGCCTCAG TATGCTGGCT CTGGAAACCT CACCCTGGCC	900
CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT	950
GAGAGCCACT CAGCTCCAGA AAAATTTGAC CTGTGAGGTG TGGGGACCCA	1000
CCTCCCCTAA GCTGATGCTG AGCTTGAAAC TGGAGAACAA GGAGGCAAAG	1050
GTCTCGAAGC GGGAGAAGCC GGTGTGGGTG CTGAACCCTG AGGCGGGGAT	1100
GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCCTGCTG GAATCCAACA	1150
TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCCAAACTC	1200
TGCTACTTGC TAGATGGAAT CCTCTTCATC TACGGAGTCA TCATCACAGC	1250
CCTGTACCTG AGAGCAAAAT TCAGCAGGAG TGCAGAGACT GCTGCCAACC	1300
TGCAGGACCC CAACCAGCTC TACAATGAGC TCAATCTAGG GCGAAGAGAG	1350
GAATATGACG TCTTGGAGAA GAAGCGGGCT CGGGATCCAG AGATGGGAGG	1400
CAAACAGCAG AGGAGGAGGA ACCCCCAGGA AGGCGTATAC AATGCACTGC	1450
AGAAAGACAA GATGCCAGAA GCCTACAGTG AGATCGGCAC AAAAGGCGAG	1500

AGGCGGAGAG	GCAAGGGGCA	CGATGGCCTT	TACCAGGACA	GCCACTTCCA 1	1550
AGCAGTGCAG	TTCGGGAACA	GAAGAGAGAG	AGAAGGTTCA	GAACTCACAA 1	1600
GGACCCTTGG	GTTAAGAGCC	CGCCCCAAAG	GTGAAAGCAC	CCAGCAGAGT	1650
AGCCAATCCT	GTGCCAGCGT	CTTCAGCATC	CCCACTCTGT	GGAGTCCATG	1700
GCCACCCAGT	AGCAGCTCCC	AGCTCTAA		:	1728

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1389 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: double

  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGAACCGGG	GAGTCCCTTT	TAGGCACTTG (	CTTCTGGTGC	TGCAACTGGC	50
GCTCCTCCCA	GCAGCCACTC	AGGGAAACAA	AGTGGTGCTG	GGCAAAAAAG	100
GGGATACAGT	GGAACTGACC	TGTACAGCTT	CCCAGAAGAA	GAGCATACAA	150
TTCCACTGGA	AAAACTCCAA	CCAGATAAAG	attctggg <b>aa</b>	ATCAGGGCTC	200
CTTCTTAACT	AAAGGTCCAT	CCAAGCTGAA	TGATCGCGCT	GACTCAAGAA	250
GAAGCCTTTG	GGACCAAGGA	AACTTCCCCC	TGATCATCAA	GAATCTTAAG	300
ATAGAAGACT	CAGATACTTA	CATCTGTGAA	GTGGAGGACC	AGAAGGAGGA	350
GGTGCAATTG	CTAGTGTTCG	GATTGACTGC	CAACTCTGAC	ACCCACCTGC	400
TTCAGGGGC	GAGCCTGACC	CTGACCTTGG	AGAGCCCCC	TGGTAGTAGC	450
CCCTCAGTG	: AATGTAGGAG	TCCAAGGGGT	AAAAACATAC	AGGGGGGAA	500
GACCCTCTC	C GTGTCTCAGC	TGGAGCTCCA	GGATAGTGGC	ACCTGGACAT	550
GCACTGTCT	C GCAGAACCAG	AAGAAGGTGG	AGTTCAAAAT	AGACATCGTG	600
GTGCTAGCT	r tccagaaggc	CTCCAGCATA	GTCTATAAGA	AAGAGGGGGA	650
ACAGGTGGA	G TTCTCCTTCC	CACTCGCCTT	TACAGTTGAA	AAGCTGACGG	700
GCAGTGGCG	A GCTGTGGTGG	CAGGCGGAGA	GGGCTTCCTC	CTCCAAGTCT	750
TGGATCACC	T TTGACCTGAI	GAACAAGGAA	GTGTCTGTA	AACGGGTTAC	800
CCAGGACCC	T AAGCTCCAGA	A TGGGCAAGAA	GCTCCCGCT	CACCTCACCC	850
TGCCCCAGG	C CITGCCTCA	TATGCTGGCT	CTGGAAACC	CACCCTGGCC	900
CTTGAAGCG	A AAACAGGAA	A GTTGCATCAG	GAAGTGAAC	C TGGTGGTGAT	950

GAGAGCCACT CAGCTCCAGA AAAATTTGAC CTGTGAGGTG TGGGGACCCA	1000
CCTCCCCTAA GCTGATGCTG AGCTTGAAAC TGGAGAACAA GGAGGCAAAG	1050
GTCTCGAAGC GGGAGAAGCC GGTGTGGGTG CTGAACCCTG AGGCGGGGAT	1100
GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCCTGCTG GAATCCAACA	1150
TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCGCAGCTC	1200
TGCTATATCC TGGATGCCAT CCTGTTTTTG TATGGTATTG TCCTTACCCT	1250
GCTCTACTGT CGACTCAAGA TCCAGGTCCG AAAGGCAGAC ATAGCCAGCC	1300
GTGAGAAATC AGATGCTGTC TACACGGGCC TGAACACCCG GAACCAGGAG	1350
ACATATGAGA CTCTGAAACA TGAGAAACCA CCCCAATAG	1389

# (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1599 base pairs

  - (B) TYPE: nucleic acid
    (C) STRANDEDNESS: double
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAACCGGG	GAGTCCCTTT	TAGGCACTIG	CITCTGGTGC	TGCAACTGGC		50
GCTCCTCCCA	GCAGCCACTC	AGGGAAACAA	AGTGGTGCTG	GGCAAAAAAG	i	100
GGGATACAGT	GGAACTGACC	TGTACAGCTT	CCCAGAAGAA	GAGCATACAA	1	L50
TTCCACTGGA	AAAACTCCAA	CCAGATAAAG	ATTCTGGGAA	ATCAGGGCTC	:	200
CTTCTTAACT	AAAGGTCCAT	CCAAGCTGAA	TGATCGCGCT	GACTCAAGAA	:	250
GAAGCCTTTG	GGACCAAGGA	AACTTCCCCC	TGATCATCAA	GAATCTTAAG		300
ATAGAAGACT	CAGATACTTA	CATCTGTGAA	GTGGAGGACC	AGAAGGAGGA		350
GGTGCAATTG	CTAGTGTTCG	GATTGACTGC	CAACTCTGAC	ACCCACCTGC		400
TTCAGGGGCA	GAGCCTGACC	CTGACCTTGG	AGAGCCCCC	TGGTAGTAGC		450
CCCTCAGTGC	AATGTAGGAG	TCCAAGGGGI	' AAAAACATAC	AGGGGGGAA		500
GACCETCTCC	GTGTCTCAGC	TGGAGCTCCA	GGATAGTGGC	ACCTGGACAT		550
GCACTGTCT	GCAGAACCAG	AAGAAGGTGG	AGTTCAAAAT	AGACATCGTG		600
GTGCTAGCT	TCCAGAAGGC	CTCCAGCATI	A GTCTATAAGA	AAGAGGGGGA		650
ACAGGTGGA	TTCTCCTTC	CACTCGCCT	TACAGTTGA	AAGCTGACGG		700
GCAGTGGCG	A GCTGTGGTG	CAGGCGGAG	A GGGCTTCCT	CTCCAAGTCT		750

TGGATCACCT	TTGACCTGAA	GAACAAGGAA	GTGTCTGTAA	AACGGGTTAC	800
CCAGGACCCT	AAGCTCCAGA	TGGGCAAGAA	GCTCCCGCTC	CACCTCACCC	850
TGCCCCAGGC	CTTGCCTCAG	TATGCTGGCT	CTGGAAACCT	CACCCTGGCC	900
CTTGAAGCGA	AAACAGGAAA	GTTGCATCAG	GAAGTGAACC	TGGTGGTGAT	950
GAGAGCCACT	CAGCTCCAGA	AAAATITGAC	CTGTGAGGTG	TGGGGACCCA	1000
CCTCCCCTAA	GCTGATGCTG	AGCTTGAAAC	TGGAGAACAA	GGAGGCAAAG	1050
GTCTCGAAGC	GGGAGAAGCC	GGTGTGGGTG	CTGAACCCTG	.AGGCGGGGAT	1100
GTGGCAGTGT	CIGCIGAGIG	ACTCGGGACA	GGTCCTGCTG	GAATCCAACA	1150
TCAAGGTTCT	GCCCACATGG	TCCACCCCGG	TGCACGCGGA	TCCCAAACTC	1200
TGCTACCTGC	TGGATGGAAT	CCTCTTCATC	TATGGTGTCA	TTCTCACTGC	1250
CITCITCCIC	AGAGTGAAGI	TCAGCAGGAG	CGCAGAGCCC	CCCGCGTACC	1300
AGCAGGGCC	GAACCAGCTC	TATAACGAGO	CTCAATCTAGG	ACGAAGAGAG	1350
GAGTACGAT	TTTTGGACA	GAGACGTGGC	CGGGACCCTG	AGATGGGGG	1400
AAAGCCGAG	A AGGAAGAAC	CTCAGGAÀGG	CCTGTACAA1	GAACTGCAGA	1450
AAGATAAGA	I GGCGGAGGC	TACAGTGAG	A TTGGGATGAI	AGGCGAGCGC	1500
CGGAGGGGC	A AGGGGCACGI	A TGGCCTTTA	C CAGGGTCTC	A GTACAGCCAC	155
CAAGGACAC	C TACGACGCC	TTCACATGC	A GGCCCTGCC	CCTCGCTAR	159

## (2) INFORMATION POR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 575 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu 10 Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu Gly Lys 25 20 Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser 40 35 Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn 55 Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala 70 Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile 85 90 Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu 105 100 Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn 115 125

Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu 135 Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly 155 Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu 165 170 Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys 185 Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser 200 Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro 215 Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp 220 230 235 Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp Leu 245 250 Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro Lys Leu 265 Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro Gln Ala Leu 280 Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala Leu Glu Ala Lys 295 Thr Gly Lys Leu His Gln Glu Val Asn Leu Val Val Met Arg Ala Thr 310 315 Gln Leu Gln Lys Asn Leu Thr Cys Glu Val Trp Gly Pro Thr Ser Pro 325 330 Lys Leu Met Leu Ser Leu Lys Leu Glu Asn Lys Glu Ala Lys Val Ser 340 345 Lys Arg Glu Lys Pro Val Trp Val Leu Asn Pro Glu Ala Gly Net Trp 360 365 Gln Cys Leu Leu Ser Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile 375 Lys Val Leu Pro Thr Trp Ser Thr Pro Val His Ala Asp Pro Lys Leu 380 390 Cys Tyr Leu Leu Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Ile Thr 395 405 410 Ala Leu Tyr Leu Arg Ala Lys Phe Ser Arg Ser Ala Glu Thr Ala Ala 420 425 Asn Leu Gln Asp Pro Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg 440 Arg Glu Glu Tyr Asp Val Leu Glu Lys Lys Arg Ala Arg Asp Pro Glu 455 Met Gly Gly Lys Gln Gln Arg Arg Arg Asn Pro Gln Glu Gly Val Tyr 470 Asn Ala Leu Gln Lys Asp Lys Met Pro Glu Ala Tyr Ser Glu Ile Gly 475 490 Thr Lys Gly Glu Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln 500 505 Asp Ser His Phe Gln Ala Val Gln Phe Gly Asn Arg Arg Glu Arg Glu 520 Gly Ser Glu Leu Thr Arg Thr Leu Gly Leu Arg Ala Arg Pro Lys Gly 535 Glu Ser Thr Gin Gln Ser Ser Gln Ser Cys Ala Ser Val Phe Ser Ile 550 565 Pro Thr Leu Trp Ser Pro Trp Pro Pro Ser Ser Ser Gln Leu 565 570

# (2) INFORMATION FOR SEQ ID NO:5:

# (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 462 amino acids

- (B) TYPE: amino acid(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gin Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu Gin Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gin Asn Gin Lys Lys Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp Gin Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp Leu Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala Leu Glu Ala Lys Thr Gly Lys Leu Eis Gln Glu Val Asn Leu Val Val Met Arg Ala Thr Gln Leu Gln Lys Asn Leu Thr Cys Glu Val Trp Gly Pro Thr Ser Pro Lys Leu Met Leu Ser Leu Lys Leu Glu Asn Lys Glu Ala Lys Val Ser Lys Arg Glu Lys Pro Val Trp Val Leu Asn Pro Glu Ala Gly Met Trp Gln Cys Leu Leu Ser Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile Lys Val Leu Pro Thr Trp Ser Thr Pro Val His Ala Asp Pro Gln Leu Cys Tyr Ile Leu Asp Ala Ile Leu Phe Leu Tyr Gly Ile Val Leu Thr Leu Leu Tyr Cys Arg Leu Lys Ile Gln Val Arg Lys Ala Asp Ile Ala Ser Arg Glu Lys Ser Asp Ala Val Tyr Thr Gly Leu Asn Thr Arg Asn 435 440 445
Gln Glu Thr Tyr Glu Thr Leu Lys His Glu Lys Pro Pro Gln
450 455 460 462

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 532 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Val Leu Gln Leu Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu Gly Lys 25 Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser 35 . 40 Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn 50 55 Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala 70 Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile 85 Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu 100 105 110 Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn \\_\_\_ 120 115 125 Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu 130 135 140 Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly 150 155 Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu 165 170 Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys 185 180 190 Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser 200 195 205 Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro 210 215 Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp 230 235 Gin Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp Leu 245 250 Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro Lys Leu 260 265 270 Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro Gln Ala Leu 280 275 285 Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala Leu Glu Ala Lys 295 300 Thr Gly Lys Leu His Gln Glu Val Asn Leu Val Val Met Arg Ala Thr 310 315 Gln Leu Gln Lys Asn Leu Thr Cys Glu Val Trp Gly Pro Thr Ser Pro 330 325 Lys Leu Met Leu Ser Leu Lys Leu Glu Asn Lys Glu Ala Lys Val Ser 345 350 Lys Arg Glu Lys Pro Val Trp Val Leu Asn Pro Glu Ala Gly Met Trp 360

Gln Cys Leu Leu Ser Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile 370 375 380 Lys Val Leu Pro Thr Trp Ser Thr Pro Val His Ala Asp Pro Lys Leu 390 395 Cys Tyr Leu Leu Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Leu Thr 410 405 Ala Leu Phe Leu Arg Val Lys Phe Ser Arg Ser Ala Glu Pro Pro Ala 420 425 430 Tyr Gln Gln Gly Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg 435 440 445 Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu 450 460 455 Met Gly Gly Lys Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn 470 475 Glu Leu Gln Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met 485 490 495 Lys Gly Glu Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly 500 505 Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Wet Gln Ala 515 Leu Pro Pro Arg 530

- (2) INFORMATION FOR SEQ ID NO:7:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 33 bases
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: nucleic acid
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

#### CGCGGGGTGA CCGTGCCCTC CAGCAGCTTG GGC

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 50 bases
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: nucleic acid
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

### CGCGGGGATC CGTCGTCCAG AGCCCGTCCA GCTCCCCGTC CTGGGCCTCA

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 bases
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

33

50

(D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: nucleic acid	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:	
CGCGGGCGGC CGCGACGCCG GCCAAGACAG CAC	33
(2) INFORMATION FOR SEQ ID NO:10:	
(i) SEQUENCE CHARACTERISTICS:  (A) LEMGTH: 33 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: nucleic acid	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:	
CGCGTTGACG AGCAGCCAGT TGGGCAGCAG CAG	33
(2) INFORMATION FOR SEQ ID NO:11:	
(i) SEQUENCE CHARACTERISTICS:  (A) LEEGTH: 15 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: nucleic acid	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:	
CGCGGGCGGC CGCTA	15
(2) INFORMATION FOR SEQ ID NO:12:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 42 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: nucleic acid	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:	
CGCGGGCTCG TTATAGAGCT GGTTCTGGCG CTGCTTCTTC TG	42
(2) INFORMATION FOR SEQ ID NO:13:	
(i) SEQUENCE CHARACTERISTICS:	

(A) LENGTH: 48 bases (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) NOLECULE TYPE: nucleic acid	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:	
CGCGGGGAGC TCGTTATAGA GCTGGTTTGC CGCCGAATTC TTATCCCG	48
(2) INFORMATION FOR SEQ ID NO:14:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 33 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: nucleic acid	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:	
CGCGGGGCGG CCACGCGTCC TCGCCAGCAC ACA	33
(2) INFORMATION FOR SEQ ID NO:15:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 36 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: nucleic acid	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:	
CGCGGGACGC GTTTCAGCCG TCCTCGCCAG CACACA	36
(2) INFORMATION FOR SEQ ID NO:16:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 33 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: nucleic acid	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:	
CGCGGGACGC GTGACCCTGA GATGGGGGGA AAG	33
(2) INFORMATION FOR SEC ID NO:17.	

	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 33 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: nucleic acid	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:	
	CGCGGGACGC GTATTGGGAT GAAAGGCGAG CGC	33
	(2) INFORMATION FOR SEQ ID NO:18:	
•	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 26 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: nucleic acid	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:	
	CCCGGATCCC AGCATGGGCA GCTCTT	26
	(2) INFORMATION FOR SEQ ID NO:19:	
	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 42 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: nucleic acid	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:	
	CGCGGGGCGG CCGCTTTAGT TATTACTGTT GACATGGTCG TT	42
	(2) INFORMATION FOR SEQ ID NO:20:	
	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 30 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: nucleic acid	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:	
	GCGGGGGGAT CCCACTGTCC AAGCTCCCAG	30
	(2) INFORMATION FOR SEQ ID NO:21:	

	(i) SEQUENCE CHARACTERISTICS (A) LENGTH: 32 bases (B) TYPE: nucleic acid (C) STRANDEDNESS: singl (D) TOPOLOGY: linear		
	(ii) MOLECULE TYPE: nucleic a	ncid	
	(xi) SEQUENCE DESCRIPTION: SE	EQ ID NO:21:	
	GCGGGGGCGG CCGCCTAAAT ACGGTTCTGG	TC 32	2
	(2) INFORMATION FOR SEQ ID NO: 22:	:	
	(i) SEQUENCE CHARACTERISTICS (A) LENGTH: 31 bases (B) TYPE: nucleic acid (C) STRANDEDNESS: sing: (D) TOPOLOGY: linear		
	(ii) MOLECULE TYPE: nucleic	acid	
	(xi) SEQUENCE DESCRIPTION: ST	EQ ID NO:22:	
٠	TCAGAAAGAG ACAACCTGAA GAAACCAACA	λ 3	1
	(2) INFORMATION FOR SEQ ID NO:23	:	
	(i) SEQUENCE CHARACTERISTIC  (A) LEEGTH: 31 bases  (B) TYPE: nucleic acid  (C) STRANDEDNESS: sing  (D) TOPOLOGY: linear	l	
	(ii) MOLECULE TYPE: nucleic	acid	

31

# (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

## TIGITGGITT CTTCAGGITG TGTCTTTCTG A

# (2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 171 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: amino acid
- (xi) SEQUENCE DESCRIPTION: SEQ ID:NO:24:

Met Glu His Ser Thr Phe Leu Ser Gly Leu Val Leu Ala Thr Leu Leu 10 Ser Gln Val Ser Pro Phe Lys Ile Pro Ile Glu Glu Leu Glu Asp Arg Val Phe Val Asn Cys Asn Thr Ser Ile Thr Trp Val Glu Gly Thr Val 40 Gly Thr Leu Leu Ser Asp Ile Thr Arg Leu Asp Leu Gly Lys Arg Ile 55 Leu Asp Pro Arg Gly Ile Tyr Arg Cys Asn Gly Thr Asp Ile Tyr Lys 70 75 Asp Lys Glu Ser Thr Val Gln Val His Tyr Arg Met Cys Gln Ser Cys 85 90 Val Glu Leu Asp Pro Ala Thr Val Ala Gly Ile Ile Val Thr Asp Val 100 105 110 Ala Ile Thr Leu Leu Leu Ala Leu Gly Val Phe Cys Phe Ala Gly His 120 125 Glu Thr Gly Arg Leu Ser Gly Ala Ala Asp Thr Gln Ala Leu Leu Arg 135 140 Asn Asp Gln Val Tyr Gln Pro Leu Arg Asp Arg Asp Asp Ala Gln Tyr 150 155 Ser His Leu Gly Gly Asn Trp Ala Arg Asn Lys 165

## (2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 182 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: amino acid
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

 Met Glu Glu Gly Lys Gly Leu Ala Val Leu Ile Leu Ala Ile Ile Leu

 5
 10
 15

 Leu Gln Gly Thr Leu Ala Gln Ser Ile Lys Gly Asn His Leu Val Lys
 20
 25
 30

 Val Tyr Asp Tyr Gln Glu Asp Gly Ser Val Leu Leu Thr Cys Asp Ala
 35
 40
 45

 Glu Ala Lys Asn Ile Thr Trp Phe Lys Asp Gly Lys Net Ile Gly Phe
 50
 60

Leu Thr Glu Asp Lys Lys Lys Trp Asn Leu Gly Ser Asn Ala Lys Asp Pro Arg Gly Met Tyr Gln Cys Lys Gly Ser Gln Asn Lys Ser Lys Pro 90 Leu Gln Val Tyr Tyr Arg Met Cys Gln Asn Cys Ile Glu Leu Asn Ala 100 105 Ala Thr Ile Ser Gly Phe Leu Phe Ala Glu Ile Val Ser Ile Phe Val 120 125 Leu Ala Val Gly Val Tyr Phe Ile Ala Gly Gln Asp Gly Val Arg Gln 130 135 140 Ser Arg Ala Ser Asp Lys Gln Thr Leu Leu Pro Asn Asp Gln Leu Tyr 150 155 Gln Pro Leu Lys Asp Arg Glu Asp Asp Gln Tyr Ser His Leu Gln Gly 165 170 Asn Gln Leu Arg Arg Asn 180

## (2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 220 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: amino acids
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Net Pro Gly Gly Leu Glu Ala Leu Arg Ala Leu Pro Leu Leu Phe 10 Leu Ser Tyr Ala Cys Leu Gly Pro Gly Cys Gln Ala Leu Arg Val Glu 20 25 Gly Gly Pro Pro Ser Leu Thr Val Asn Leu Gly Glu Glu Ala Arg Leu 40 Thr Cys Glu Asn Asn Gly Arg Asn Pro Asn Ile Thr Trp Trp Phe Ser 55 60 Leu Gln Ser Asn Ile Thr Trp Pro Pro Val Pro Leu Gly Pro Gly Gln 70 75 Gly Thr Thr Gly Gln Leu Phe Phe Pro Glu Val Asn Lys Asn Thr Gly 90 Ala Cys Thr Gly Cys Gln Val Ile Glu Asn Asn Ile Leu Lys Arg Ser 100 105 Cys Gly Thr Tyr Leu Arg Val Arg Asn Pro Val Pro Arg Pro Phe Leu 115 120 Asp Met Gly Glu Gly Thr Lys Asm Arg Ile Ile Thr Ala Glu Gly Ile 135 140 Ile Leu Leu Phe Cys Ala Val Val Pro Gly Thr Leu Leu Leu Phe Arg 150 155 Lys Arg Trp Gln Asn Glu Lys Phe Gly Val Asp Met Pro Asp Asp Tyr 170 175 Glu Asp Glu Asn Leu Tyr Glu Gly Leu Asn Leu Asp Asp Cys Ser Met 180 185 190 Tyr Glu Asp Ile Ser Arg Gly Leu Gln Gly Thr Tyr Gln Asp Val Gly 200 Asn Leu His Ile Gly Asp Ala Gln Leu Glu Lys Pro 210

(2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 228 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: amino acid
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met Ala Thr Leu Val Leu Ser Ser Met Pro Cys His Trp Leu Leu Phe Leu Leu Leu Phe Ser Gly Glu Pro Val Pro Ala Met Thr Ser Ser 20 25 Asp Leu Pro Leu Asn Phe Gln Gly Ser Pro Cys Ser Gln Ile Trp Gln 40 His Pro Arg Phe Ala Ala Lys Lys Arg Ser Ser Met Val Lys Phe His 55 Cys Tyr Thr Asn His Ser Gly Ala Leu Thr Trp Phe Arg Lys Arg Gly 70 Ser Gln Gln Pro Gln Glu Leu Val Ser Glu Glu Gly Arg Ile Val Gln 85 90 Thr Gln Asn Gly Ser Val Tyr Thr Leu Thr Ile Gln Asn Ile Gln Tyr 105 100 110 Glu Asp Asn Gly Ile Tyr Phe Cys Lys Gln Lys Cys Asp Ser Ala Asn 120 125 His Asn Val Thr Asp Ser Cys Gly Thr Glu Leu Leu Val Leu Gly Phe 135 140 Ser Thr Leu Asp Gln Leu Lys Arg Arg Asn Thr Leu Lys Asp Gly Ile 145 150 155 160 Ile Leu Ile Gln Thr Leu Leu Ile Ile Leu Phe Ile Ile Val Pro Ile 165 ~ 170 Phe Leu Leu Asp Lys Asp Asp Gly Lys Ala Gly Met Glu Glu Asp 180 185 190 His Thr Tyr Glu Gly Leu Asn Ile Asp Gln Thr Ala Thr Tyr Glu Asp 200 195 205 Ile Val Thr Leu Arg Thr Gly Glu Val Lys Trp Ser Val Gly Glu His 210 215 220 Pro Gly Gln Glu 225